

Online Research @ Cardiff

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository: <https://orca.cardiff.ac.uk/id/eprint/59375/>

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Palma, Leopoldo, Munoz, Delia, Berry, Colin ORCID: <https://orcid.org/0000-0002-9943-548X>, Murillo, Jesus and Caballero, Primitivo 2014. Draft genome sequences of two *Bacillus thuringiensis* strains and characterization of a putative 41.9-kDa insecticidal protein. *Toxins* 6 (5), pp. 1490-1504. 10.3390/toxins6051490 file

Publishers page: <http://dx.doi.org/10.3390/toxins6051490>
<<http://dx.doi.org/10.3390/toxins6051490>>

Please note:

Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies.

See

<http://orca.cf.ac.uk/policies.html> for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.



Article

Draft Genome Sequences of Two *Bacillus thuringiensis* Strains and Characterization of a Putative 41.9-kDa Insecticidal Toxin

Leopoldo Palma ¹, Delia Muñoz ², Colin Berry ³, Jesús Murillo ² and Primitivo Caballero ^{1,2,*}

¹ Instituto de Agrobiotecnología, Consejo Superior de Investigaciones Científicas-Universidad Pública de Navarra-Gobierno de Navarra, Campus Arrosadía, Mutilva Baja, Navarra 31192, Spain; E-Mail: leopoldo.palma@unavarra.es

² Grupo de Protección Cultivos, Departamento de Producción Agraria, Escuela Técnica Superior de Ingenieros Agrónomos, Universidad Pública de Navarra, Pamplona, Navarra 31006, Spain; E-Mails: dmunoz@unavarra.es (D.M.); jesus.murillo@unavarra.es (J.M.)

³ Cardiff School of Biosciences, Cardiff University, Park Place, Cardiff CF10 3AT, UK; E-Mail: Berry@cf.ac.uk

* Author to whom correspondence should be addressed; E-Mail: pcm92@unavarra.es; Tel.: +34-948-168-004; Fax: +34-948-169-300.

Received: 19 February 2014; in revised form: 21 April 2014 / Accepted: 24 April 2014 /

Published: 30 April 2014

Abstract: In this work, we report the genome sequencing of two *Bacillus thuringiensis* strains using Illumina next-generation sequencing technology (NGS). Strain Hu4-2, toxic to many lepidopteran pest species and to some mosquitoes, encoded genes for two insecticidal crystal (Cry) proteins, *cryIIa* and *cry9Ea*, and a vegetative insecticidal protein (Vip) gene, *vip3Ca2*. Strain Leapi01 contained genes coding for seven Cry proteins (*cryIAa*, *cryICa*, *cryIDa*, *cry2Ab*, *cry9Ea* and two *cryIIa* gene variants) and a *vip3* gene (*vip3Aa10*). A putative novel insecticidal protein gene 1143 bp long was found in both strains, whose sequences exhibited 100% nucleotide identity. The predicted protein showed 57 and 100% pairwise identity to protein sequence 72 from a patented Bt strain (US8318900) and to a putative 41.9-kDa insecticidal toxin from *Bacillus cereus*, respectively. The 41.9-kDa protein, containing a C-terminal 6× HisTag fusion, was expressed in *Escherichia coli* and tested for the first time against four lepidopteran species (*Mamestra brassicae*, *Ostrinia nubilalis*, *Spodoptera frugiperda* and *S. littoralis*) and the green-peach aphid *Myzus persicae* at doses as high as 4.8 µg/cm² and 1.5 mg/mL, respectively. At these protein concentrations, the recombinant 41.9-kDa protein caused no mortality or symptoms of impaired growth against any of the insects tested, suggesting that

these species are outside the protein's target range or that the protein may not, in fact, be toxic. While the use of the polymerase chain reaction has allowed a significant increase in the number of Bt insecticidal genes characterized to date, novel NGS technologies promise a much faster, cheaper and efficient screening of Bt pesticidal proteins.

Keywords: *Bacillus thuringiensis*; insecticidal toxins; next-generation sequencing; genome annotation; microbial control; insecticidal activity

1. Introduction

Bacillus thuringiensis (Bt) is one of the best-characterized entomopathogenic bacteria, with many strains bearing a wide variety of insecticidal genes [1]. Delta-endotoxins (Cry and Cyt proteins), synthesized during the stationary growth phase as crystalline, parasporal inclusions, are highly active against a wide range of insects [1,2] but also against nematodes [3] and a human-pathogenic protozoan [4]. Parasporins (Ps), another crystal protein group of Bt, may possess the typical three-domains of Cry proteins (Ps1 and Ps3 but not Ps2 or Ps4) and, although lacking insecticidal activity, exhibit strong cytotoxic activity against human cancer cells of various origins upon protease activation [5,6]. Crystal proteins accumulate in the cells as parasporal inclusions and can account for up to 25% of the sporulated cell dry weight. Under laboratory conditions, Bt is able to produce nearly 0.5 mg of crystal proteins per mL [7]. In order to produce this large amount of protein and form crystals, some insecticidal proteins require the presence of helper proteins, *i.e.*, P19 and P20 for the stable production of Cyt1A and Cry11A [8]. P20 also enhances the expression and crystallization of Cry1Ac in acrySTALLIFEROUS and plasmid-negative Bt strains [9] and synergizes the toxicity of Cry11A against third-instar larvae of *Aedes aegypti* (Diptera: Culicidae) [10]. Bt synthesizes additional insecticidal proteins during the vegetative growth phase, which are subsequently secreted into the growth medium. These proteins are commonly known as vegetative insecticidal proteins (Vips) and hold insecticidal activity against lepidopteran [11], coleopteran [12] and some homopteran pests [13]. A less characterized secretory protein with no amino acid similarity to previously known Vips, and termed Sip protein (secreted insecticidal protein), has insecticidal activity against coleopteran pests [14]. The only Sip protein discovered to date is a 41-kDa protein, with high toxicity against *Leptinotarsa decemlineata* (Coleoptera: Chrysomelidae) and inhibits larval development of *Diabrotica undecimpunctata howardi* and *Diabrotica virgifera virgifera* (Coleoptera: Chrysomelidae) [14]. Bt also produces Mtx-like and Bin-like proteins, which share amino acid similarities with mosquitocidal Mtx and Bin toxins, respectively, from *Lysinibacillus sphaericus* (formerly *Bacillus sphaericus*) [15,16]. In summary, Bt insecticidal proteins may be classified into at least five distinct protein groups according to their amino acid identity and protein structure: Bin-like, three-domain Cry toxins, Cyt, Mtx-like, Sip, and Vip proteins [16].

In addition to the insecticidal toxins mentioned above, some *B. thuringiensis* strains produce other virulence factors such as phospholipases, proteases, haemolysins, enterotoxins, enhancin-like proteins (metalloproteases) and chitinases, plus extracellular (secreted) antimicrobial compounds (AMPs) such as bacteriocins and β -exotoxins, which may contribute to the pathogenicity of this bacterium against

insects [1]. The phospholipases, haemolysins, enhancin-like proteins and chitinases are synthesized in the insect midgut during vegetative growth, facilitating the dissemination of the bacteria through the peritrophic membrane and towards the hemocoel [1,17]. The β -exotoxins, also known as thuringiensins, are secreted, heat-stable, secondary metabolites analogues to the nucleotide adenine and with low molecular weight (~700 Da) [18].

Screening programs have identified thousands of Bt strains and insecticidal genes active against a wide range of insect orders and several nematodes, mites and protozoans [16]. However, such screening programs have been usually costly and lengthy and not always efficient. Generally, toxin genes are identified by polymerase chain reaction (PCR) analyses, along with restriction fragment length polymorphism (RFLP) analysis of the amplicons, followed by time-consuming chromosome-walking techniques to recover the full-length coding sequence of each toxin [19–21]. Today, next-generation sequencing technologies (NGS) allow rapid sequencing of entire genomes with profitable cost-effective ratios [22,23]. This study aimed to sequence and analyze the genome of two Bt strains containing known insecticidal genes using Illumina NGS-sequencing technology and to characterize a putative 41.9-kDa insecticidal protein of unreported biological activity.

2. Materials and Methods

2.1. Bacterial Strains and Plasmids

The Bt strains used in this work belong to a collection maintained at the Universidad Pública de Navarra (Pamplona, Spain) which comprises more than 2000 strains organized in two major sub-collections: one from mainland Spain and a second one from the Canary Islands [24,25]. Strains Hu4-2 and Leapi01 are from the Spanish mainland sub-collection. Strain Leapi01 was isolated from a dead *Mythimna loreyi* (Lepidoptera: Noctuidae) larva during a natural epizootic in corn crops in the province of Badajoz (Spain) [26]. Strain Hu4-2 was isolated from dust samples obtained in a maize grain silo from the province of Huesca (Spain) [27]. Bt strains and *E. coli* DH5 α , used for routine DNA manipulations, were cultured in Luria-Bertani (LB) medium (1% tryptone, 0.5% yeast extract, and 1% NaCl, pH 7.0) at 28 °C and 37 °C, respectively. Strain *E. coli* BL21(DE3) was used for protein expression and was propagated in 2 \times YT medium (1.6% tryptone, 1% yeast extract, and 0.5% NaCl, pH 7.0) at 37 °C. Vectors pGEM-T Easy (Promega) and pET-28b(+) (Novagen) were used for routine cloning of PCR products and for the production of recombinant protein, respectively.

2.2. DNA Isolation, Sequencing and Computational Analysis

For genome sequencing, total DNA (chromosome and plasmids) was isolated using the Wizard Genomic DNA Purification Kit (Promega) following the instructions for DNA isolation from Gram-positive bacteria. Purified total DNA from strain Hu4-2 was used to construct a pooled Illumina library and sequenced on a HiSeq 2000 Sequencing System (Illumina Sequencing) in a single read mode with a read length of 50 bases (GATC Biotech, Konstanz, Germany). Reads were then assembled using the CLC Genomic Workbench software (CLC Bio, Aarhus, Denmark) with the *de novo* assembly tool and default parameters. Purified total DNA from Leapi01 was sequenced at the Beijing Genomics Institute (BGI, Shenzhen, China) using high-throughput Illumina sequencing technology

and a paired-end library sequencing of 500 bp. Reads were assembled using SOAP denovo (version 1.05) with an optimized k-mer size of 59-mer. Identification and annotation of insecticidal genes was done with BLAST [28] using custom insecticidal toxin databases and the BtToxin_scanner [29], and complemented with the NCBI Prokaryotic Genome Annotation Pipeline (released 2013) and the RAST server [30]. The custom insecticidal toxin database was constructed with amino acid sequences of Cry and Cyt proteins [1], mosquitocidal proteins Mtx1 and Mtx2 [15], the secreted insecticidal protein Sip [14], vegetative insecticidal proteins Vip1/Vip2 and Vip3 [11,12], cancer cell-killing Cry proteins or parasporins [5,6], accessory proteins P19 and P20 [8] and the haemolytic enterotoxin HBL from *Bacillus cereus* [31]. Multiple sequence alignments, signal peptide and conserved domain analyses plus primer design were performed using suitable tools included in Geneious Pro v6.1.4 [32].

2.3. Amplification and Cloning of the Full-Length 41.9-kDa-Protein Gene

Primers were designed to amplify the full-length CDS sequence of the putative 41.9-kDa protein but excluding the stop codon, allowing the 6× HisTag fusion of the pET-28b(+) vector into the 3' terminus of the gene. Forward and reverse primers included a 5'-*Nco*I and a 3'-*Sal*I restriction sites, respectively. The PCR reaction was performed in a 25 µL reaction mixture containing 5 µL 5× reaction buffer, 10 mM dNTP mixture, 6.3 pmol each of forward and reverse primers, 0.5 U proof reading PrimeSTAR HS DNA polymerase (Takara) and 100 ng total DNA in an C1000 Touch thermal cycler (Bio-Rad) using the following cycling conditions: 4 min initial denaturation at 94 °C, 35 amplification cycles (1 min denaturation at 94 °C, 1 min annealing at 52 °C and 72 s extension at 72 °C) and a final extension step at 72 °C for 10 min. Electrophoresed PCR products (1% agarose gels) were gel-purified using NucleoSpin Extract II kit (Macherey-Nagel), end-modified by an A-tailing procedure and ligated into pGEM-T easy according to the manufacturer's instructions (Promega). Ligation mixtures were then transformed into *E. coli* DH5α using standard procedures [33]. Plasmid DNA was purified using the NucleoSpin Plasmid kit (Macherey-Nagel) and sequenced (Sistemas Genómicos, Valencia, Spain). A clone containing an insert identical to the genome sequence of the 41.9 kDa protein CDS was selected, the insert was excised by digestion with *Nco*I and *Sal*I, purified from the agarose gel, ligated into predigested pET-28b(+) and transformed into *E. coli* BL21(DE3).

2.4. Expression and Purification of the 41.9-kDa Recombinant Protein

An *E. coli* BL21(DE3) clone harboring the pET28-b(+) vector containing the 41.9-kDa protein CDS sequence was pre-cultured overnight at 37 °C and 200 rpm in 2× YT medium containing 50 µg/mL kanamycin. A 1/25 dilution of this pre-culture was inoculated into 1 liter 2× YT medium containing 50 µg/mL kanamycin and further incubated with vigorous shaking (250 rpm) until OD₆₀₀ = 0.7–1.0. Expression was induced immediately with 1mM isopropyl-β-D-1-thiogalactopyranoside (IPTG, final concentration) and the culture was incubated for a further 16 h at 37 °C. Samples were centrifuged at 5000 g for 15 min at 4 °C and the resulting pellet was weighed and resuspended in 3 mL sonication buffer per gram of pellet, containing 20 mM sodium phosphate buffer pH 7.4, 0.5 M NaCl, 3 mg/mL lysozyme, 25 U/µL Benzonase (Novagen) per mL of buffer and 100 µM phenylmethylsulfonyl fluoride (PMSF). Samples were further incubated at 37 °C for 1 h and sonicated on iced-water with a Branson Analog Sonifier S-250 (Fisher Scientific) by applying two 1-min pulses

with 60% constant duty cycle and separated by a 1-min cooling period. Insoluble material was removed by centrifugation at 12,000 g for 30 min at 4 °C and the soluble cellular fraction filtered through sterile 0.45 and 0.22 µm syringe filters. Protein purification was performed at room temperature (RT) using Protino Ni-TED 2000 Packed Columns according to the manufacturer's instructions (Macherey-Nagel). Once the bound polyhistidine-tagged protein was eluted, the buffer exchange procedure was performed immediately at room temperature with Milli-Q water and GE Healthcare PD-10 desalting columns to prevent protein aggregation and precipitation. This step also prevents the potential toxic effects of imidazole [34,35] to the insects tested in bioassays. The resultant expressed protein was then analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE), stained with Coomassie brilliant blue R-250 (Sigma-Aldrich). Duplicate unstained Coomassie SDS-PAGE gel was also analyzed by Western Blot using anti-His-tag antibodies (Sigma-Aldrich, Seelze, Germany) [33]. The protein concentration was quantified by the Bradford method [36].

2.5. Insect Rearing and Bioassays

For novel putative toxins of unknown biological activity, the highest protein concentrations (limited by amount of protein produced in expression experiments) were used in bioassays to detect any sign of toxicity (e.g., mortality or impaired growth) for the species tested. If toxicity was detected, we carried out bioassays using lower protein concentrations in order to establish dose-mortality responses and LC₅₀ (mean lethal concentration) values. For qualitative bioassays with *Myzus persicae* (Homoptera: Aphididae), cohorts of second instar nymphs were obtained from gravid females during a two-day period prior to the experiment. To evaluate insecticidal activity, the purified (soluble) recombinant protein was incorporated into a liquid diet containing 20% (w/v) sucrose in Milli-Q water at a single high concentration of 1.5 mg/mL. Fifteen-second instar nymphs were placed inside a cylindrical plastic cage without lid (3 cm diameter × 1.5 cm height). The cages were then sealed with a Parafilm layer, on which 100 µL drops of protein-diluted diet or protein-free diet (as negative controls) were loaded and confined using a second Parafilm layer, then aphids were allowed to feed through the Parafilm membrane. Bioassays were conducted at 25 °C, 60 ± 5% RH, and a 16:8 h light:dark photoperiod and mortality was recorded after 72 h. For other bioassays, 24 neonate larvae of four lepidopteran species, one from the Pyralidae family, *Ostrinia nubilalis*, and three from the Noctuidae family, *Mamestra brassicae*, *Spodoptera frugiperda* and *S. littoralis*, were used. The purified (soluble) recombinant protein diluted in water was poured onto each well of 24-well plates (35 µL/well) containing a layer of artificial diet and allowed to dry at a final concentration of 4.8 µg/cm² and one neonate larvae per well. Water was also used instead of protein dilutions in negative controls and the bioassay was repeated twice. Bioassays were conducted at 25 °C, 60% ± 5% RH, and a 16:8 h light:dark photoperiod and mortality was recorded after seven days.

2.6. Nucleotide Sequences Accession Numbers

Draft genome sequences have been deposited in GenBank under accessions AMXT000000000 (GenBank BioProject PRJNA175233) and AMXS000000000 (Genbank BioProject PRJNA175222) for strains Hu4-2 and Leapi01 strains, respectively.

3. Results and Discussion

Finding novel proteins is particularly important to manage the increasing resistance occurrence to Bt-based insecticides or Bt-plants reported for some species [37–39]. Today, next-generation sequencing (NGS) technologies [22,23,40–42] arise as a novel and useful tool for the discovery of completely novel insecticidal-toxin genes that would otherwise be difficult to identify. NGS technologies provides a fast and reliable framework to obtain complete genomic sequences, and offers excellent cost-effective ratios since complete genomes sequences can be obtained for less than 10^{-4} euro per nucleotide. In this work, an efficient strategy combining Illumina *de novo* sequencing and a genome annotation pipeline, which allowed the rapid identification of all kinds of insecticidal genes in sequenced Bt strains, has been performed, and a novel putative 41.9-kDa insecticidal protein analyzed.

3.1. Genome Assembly and Annotation of Bt Strains Hu4-2 and Leapi01

As many as 393 contigs were obtained from strain Hu4-2 totaling 6,458,628 bp with the N50 value [43] at 43,448 bp and 34.7% G+C content. Genome annotation produced 6888 coding sequences (CDSs) with 44 pseudogenes and 37 RNA genes (ribosomal and tRNAs). The assembly contained two complete insecticidal crystal protein genes (*cryIIa7* and *cry9Ea*), a vegetative insecticidal protein gene (*vip3Ca2*) (Table 1), which was flanked by putative transposase genes (data not shown), and a CDS coding for a protein with 62% pairwise identity to a putative 41.9-kDa insecticidal toxin from *B. cereus* (Table 1). The draft genome for strain Hu4-2 also included short incomplete sequences exhibiting homology to several Cry toxins (Table 1). We do not know whether these incomplete sequences represent true incomplete genes or are the result of misassemblies, which could be due to the existence of regions of high identity among different insecticidal genes and the short read lengths. In addition, this strain also contained two short vestigial toxin gene remnants (VTGR) or pseudogenes (Table 1), with internal stop codons and overlapping BLASTx hits in different frames (frame-shifts), exhibiting 98% and 46% pairwise identity to a *cry2Ab* and *vip3Ba* genes. Two CDSs with homology to the helper protein P19 (accession No. AAB34196) were also found exhibiting 44% and 47% pairwise identity and located upstream of incomplete (lacking the 3' end) *cryICa* and *cryIFa* sequences, respectively. The complete haemolysin (*hbl*) coding operon from *B. cereus* was also found. The Hu4-2 assembly showed no homology with other toxins such as Mtx toxins, Vip1/Vip2, Sip or parasporin proteins.

Table 1. Insecticidal proteins showing homology with the draft genomes of Bt strains Hu4-2 and Leapi01.

| Strain | Closest homolog | Accession number | % aa identity | CDS completeness |
|--------|-----------------|------------------|---------------|------------------|
| Hu4-2 | Cry1Ad | AAA22340 | 100 | Incomplete |
| | Vip3Ba | AAV70653 | 46 | VTGR |
| | Cry9Ea | BAA34908 | 100 | Full-length |
| | Cry1Ab | AAA22330 | 100 | Incomplete |
| | Cry1Ad | AAA22340 | 100 | Incomplete |
| | Cry1Fa | AAA22348 | 100 | Incomplete |

Table 1. Cont.

| Strain | Closest homolog | Accession number | % aa identity | CDS completeness |
|---------|-----------------|------------------|---------------|------------------|
| Hu4-2 | Vip3Ca2 | AEE98106 | 100 | Full-length |
| | 41.9 kDa | ZP_04231421 | 62 | Full-length |
| | Cry1Da | CAA38099 | 100 | Incomplete |
| | Cry1Ia7 | AAM73516 | 100 | Full-length |
| | Cry1Ca+P19 | CAA30396 | 100 | Incomplete |
| | Cry1Aa | AAA22353 | 100 | Incomplete |
| | Cry1Ca | CAA30396 | 100 | Incomplete |
| | Cry1Fa | AAA22348 | 100 | Incomplete |
| | Cry1Fa | AAA22348 | 100 | Incomplete |
| | Cry1Ea | CAA37933 | 99 | Incomplete |
| | Cry1Da | CAA38099 | 100 | Incomplete |
| | Cry2Ab | ACC86136 | 98 | VTGR |
| | Cry1Fa+P19 | AAA22348 | 100 | Incomplete |
| | Cry1Ab | AAA22330 | 100 | Incomplete |
| | Cry1Ad | Q03744 | 99 | Incomplete |
| | Cry1Fa | AAA22348 | 100 | Incomplete |
| Leapi01 | Cry1Aa | AAA22330 | 99 | Incomplete |
| | Cry1Ab | AAA22330 | 100 | Incomplete |
| | Cry9Ea | CAA85764 | 100 | Full-length |
| | Cry1Ca | CAA30396 | 100 | Full-length * |
| | Cry1Da | CAA38099 | 100 | Full-length * |
| | Cry1Ia2 | CAA44633 | 100 | Full-length |
| | Cry1Ab | AAA22330 | 100 | VTGR |
| | Cry1Aa | AAA22353 | 100 | Incomplete |
| | Cry1Ab | AAA22330 | 100 | Incomplete |
| | 41.9-kDa | ZP_04231421 | 62 | Full-length |
| | Vip3Aa10 | AAC37036 | 99 | Full-length |
| | Cry2Ab | AAA22342 | 100 | Full-length |
| | Cry1Aa | BAA00257 | 100 | Full-length ** |
| | Cry1Ia14 | ACG63871 | 100 | Full-length ** |

Notes: Incomplete: unfinished contig sequences; VTGR: vestigial toxin gene remnants or pseudogenes;

* Split by N-in-frame characters, the full-length coding sequence was deduced by alignment; ** The full-length coding sequence was re-constructed by re-assembly with plasmid pCT281 sequence; Acc. num.: GenBank accession number.

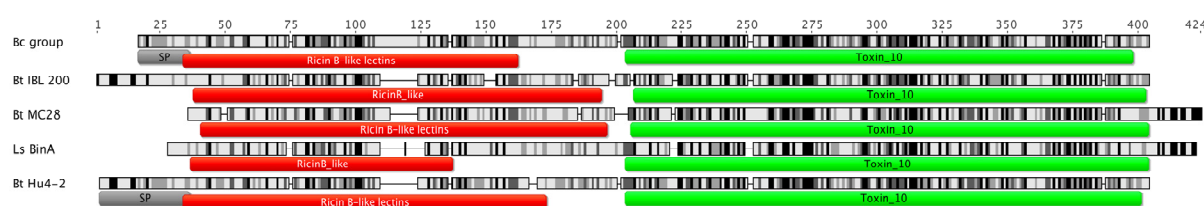
The Leapi01 draft genome produced 225 scaffolds totaling 6,156,720 bp with the N50 value at 31,921 bp and 34.8% G+C content. Genome annotation identified 6158 CDSs with 40 pseudogenes and 41 RNA genes. This strain carried seven complete insecticidal crystal protein genes (*cry1Aa*, *cry1Ca*, *cry1Da*, *cry2Ab*, *cry9Ea* and two *cry1Ia* variants) and a *vip* gene (*vip3Aa10*) that, as occurred with the *vip3Ca2* gene of strain Hu4-2, was flanked by two sequences encoding putative transposases (data not shown). This strain also contained the CDS coding for the putative 41.9-kDa insecticidal toxin found in Hu4-2 strain (Table 1) and the complete haemolysin (*hbl*) coding operon previously found in strain Hu4-2. Genes *cry1Ca*, *cry1Da* and *cry1Ia2* were located closely together into the

same scaffold, which also encoded a putative helper protein with 47.7% identity to P19 (accession No. AJ010753). Scaffold 118 (18,932 bp) exhibited 99% pairwise identity with plasmid pCT281 (281, 231 bp) from Bt strain CT-43. To evaluate the possibility that strain Leapi01 could contain a plasmid similar to pCT281, we performed a reassembly of the Leapi01 reads using pCT281 as reference sequence. This resulted in a consensus sequence of 281,256 bp showing synteny and 98.7% pairwise identity to pCT281. This consensus included genes *cryIAa3*, *cryIIa14*, *cry2Ab* and *vip3Aa10*; however, gene *cry2Aa9* was not included in this or any of the other contigs of the assembly, although it is present in plasmid pCT281. In addition to the toxin genes described above, this strain also contained short contigs with homology to diverse *cryIA* genes and a short VTGR sequence with homology to *cryIAb* gene (Table 1). This strain did not contain sequences with homology to Mtx toxins, Vip1/Vip2, Sip or parasporins. Finally, the *vip3* genes from strains Hu4-2 and Leapi01 are preceded by a CDS coding for a trans-acting positive regulator protein (Multiple Virulence Gene Regulator, Mga from group A *Streptococcus*) with 94.9% amino acid identity between both strains [44].

3.2. Molecular Characterization of the Putative 41.9-kDa Insecticidal Protein

The 41.9-kDa genes encoded by Leapi01 and Hu4-2 strains were identical and 1143 bp long, with a deduced amino acid sequence of 380 residues and predicted molecular weight of 41.9 kDa. Their predicted protein sequences exhibited 57 and 100% pairwise identity to protein sequence 72 from patent US8318900 and the 41.9-kDa component of a putative insecticidal toxin from *B. cereus*, respectively. They were identical to the hypothetical protein IE5_04900 from *B. cereus* BAG3X2-2 and showed 38 and 60% identity to BinA and BinB, respectively, the two components of the Bin binary toxin from *L. sphaericus* (Bacillales: Planococcaceae). Protein analysis with Geneious Pro detected a putative signal peptide sequence from residues 1 to 35 (residue positions of the 41.9-kDa protein sequence) plus two conserved domains: a ricin B-like lectin (InterPro family IPR000772) and a Toxin_10 conserved domain commonly found in some binary insecticidal crystal toxins (InterPro family IPR008772) from residues 33 to 153 and from 183 to 377 (residue positions of the 41.9-kDa protein sequence), respectively (Figure 1).

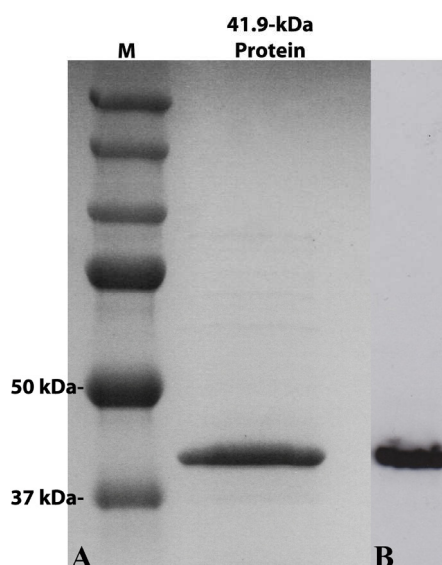
Figure 1. Comparison between *B. cereus* group (Bc) hypothetical protein homologous to 41.9-kDa proteins, *B. thuringiensis* (Bt) 41.9-kDa insecticidal protein from strains IBL 200, MC28 and Hu4-2 (identical to that of strain Leapi01), and *L. sphaericus* (Ls) BinA protein. (SP) predicted signal peptide. Sequence identity is indicated by shading: black for 100%, dark grey for 80%–100%, light grey for 60%–80%, and white for less than 60% identity. Scale indicates residue numbers in the multiple sequence alignment.



3.3. Insecticidal Bioassay of the Putative 41-kDa Insecticidal Protein

The gene for the 41.9-kDa protein from strain Hu4-2 was successfully expressed in *E. coli* and produced a major protein of the expected size (approx. 42 kDa) in a soluble form (Figure 2), as observed in SDS-PAGE analysis. At the concentrations used, the purified (soluble) recombinant 41.9-kDa protein caused no mortality or impaired growth of any of the lepidopteran insects tested, namely: *M. brassicae*, *O. nubilalis*, *S. frugiperda* and *S. littoralis*, nor of the green-peach aphid *M. persicae*.

Figure 2. SDS-PAGE gel (A) and Western Blot (B) analysis of the 41.9-kDa recombinant protein expressed in *E. coli* (right lane). The recombinant protein was tagged with a 6× His in its C-terminal end and purified from a nickel column before separation by electrophoresis; after transfer to a nylon membrane, the protein was detected using anti-His-tag antibodies. A molecular weight marker (Precision plus protein dual color standards, Bio-Rad) was electrophoresed along with the sample and the sizes of the fragments are indicated to the left of the panel.



4. Conclusions

In this work, we report the draft genome sequences of the Bt strains Hu4-2 and Leapi01, which have been previously characterized with different levels of description [27,45]. Their insecticidal gene content was explored mainly using the BtToxin_scanner [29] and custom BLAST searches [28]. The BtToxin_scanner has previously been demonstrated to be an efficient tool for the identification of insecticidal toxin coding sequences over both nucleotide and protein sequences [29]. It also detected both short vestigial and full-length toxin-coding sequences even with in-frame ambiguous (e.g., N) characters. This situation occurs when NGS data are assembled with sequencing data obtained from pair-end libraries (*i.e.*, Leapi01 strain), which allows for the assessment of the order, distance and orientation of contigs along with their re-assembly into larger sequences or scaffolds [46]. These contigs might contain gaps, which are filled with as many Ns as the most likely estimate of their length [47]. However, the inability of the scanner to detect other valuable proteins (*i.e.*, Sip, Bin,

Mtx-like toxins and accessory P19/P20 proteins) made the combination of both methods (custom BLAST searches and scanning) the best approach. Annotation data indicated that even though both strains showed different toxin repertoires, they shared three complete genes: the insecticidal crystal protein gene *cry9Ea*, the putative 41.9-kDa insecticidal protein gene and different *cryIIa* gene variants (Table 1). The two strains also showed several short sequences encoding toxin fragments (Table 1) of two types: those likely located in partially sequenced or unfinished contigs, commonly found in unclosed draft genomic sequences and short sequences of vestigial toxin gene remnants or pseudogenes, as seen in previous sequencing studies [48]. From the available data, it is not possible to conclude whether the partial sequences represent fragments of whole CDSs or pseudogenes. The pseudogenes, representing residual coding sequences, which seem to be the leftovers of genes undergoing recombination or transposition events, especially when flanked by mobile genetic elements, may constitute proof of toxin evolution and explain the distribution of genes in different strains [49]. In fact, full-length *vip3* genes (*vip3Aa10* and *vip3Ca2*) were also found flanked by short sequences (100 to 300 bp) with homology to putative transposases. In addition, upstream of these genes can be found a conserved CDS with homology to a trans-acting positive regulator protein (Mga), a DNA-binding protein that activates the expression of several important virulence genes in group-A *Streptococcus* [50]. The high similarity of contig 188 from the Leapi01 strain with plasmid pCT281 (~99%), from the Bt strain CT-43 [51], indicated the likely presence of this plasmid in Leapi01. In addition, genes *cryIAa3*, *cryIIa14* and *cry2Aa9* have been found close together in a 28,171 bp region of pCT281, constituting a pathogenicity island in strain CT-43 [51]. The other nine-plasmid sequences described for this strain were also screened in Leapi01 or Hu4-2, but they were not detected. Most of the insecticidal genes present in both strains (except that coding for the 41.9-kDa protein) were located in contigs or scaffolds with the highest levels of read coverage and without mapping over reference Bt genomes, indicating they are probably present in high-copy plasmids. However, our work was focused primarily in finding full-length insecticidal genes for cloning independently of their location (plasmid or chromosome). Bt has been considered a safe and environmentally friendly bacterium for the control and management of a wide number of the most damaging insect pests worldwide because of its lack of toxicity and pathogenicity against mammals [52]. However, some strains (*i.e.*, Bt strain 97-27) carry chromosome-encoded virulence factors commonly found in the mammalian opportunistic pathogen, *B. cereus*, that threaten its safety record [53]. Bt is closely related to *B. cereus* and *B. anthracis* [54], and strain 97-27 bears the *hbl* operon, with haemolytic enterotoxin genes *hblCDBA*, which is suspected to cause diarrhea by *B. cereus* contaminated food. Nevertheless, some strains PCR-positive for the *hblA* gene (*i.e.*, Bt strain WS2623) are not cytotoxic or haemolytic, suggesting that this gene is present but incomplete or silent, whereas others, despite their demonstrated safe use as biopesticide, produced the enterotoxin HBL [31]. In this work, both strains exhibited the complete homologous *hbl* operon, but they have not been tested yet for haemolytic activity. The predicted 41.9-kDa protein resembles previously known insecticidal toxins and shares 38 and 60% maximum pairwise identity to components A and B of Bin binary toxin from *L. sphaericus*, respectively. This *L. sphaericus* binary toxin is highly toxic against mosquitoes [15]. Both components are co-transcribed from a single operon before the end of the exponential growth and sporulation as a crystal toxin [15]. The Bin protein sequences are extremely conserved between strains and only five variants have been reported to date with a few differences in their amino acid sequences [55]. In contrast, the 41.9-kDa-protein

gene from the Spanish strains was found as a single coding sequence and its deduced amino acid sequence was poorly conserved among those of different Bt strains (34% maximum pairwise identity), in which 41.9-kDa related proteins were also found as single coding sequences. These findings strongly suggested that this protein might have a non-binary behavior and work as a single toxin. When compared with the *L. sphaericus* BinB protein, the 41.9 kDa protein shared the two conserved domains but not the predicted signal peptide and the transmembrane domain, which were found exclusively in the 41.9-kDa protein from the Spanish strains (Figure 1) and suggest that it might be secreted. In fact, this protein has not been detected in previous SDS-PAGE analyses of Leapi01 and Hu4-2 crystal proteins, consistent with synthesis during the vegetative growth phase and subsequent secretion into the growth medium [26,27]. A similar situation occurs with other secreted insecticidal proteins such as Vip3 [11] and CryIIa7 proteins [56]. To our knowledge, this study represents the first bioassay of the 41.9-kDa putative insecticidal proteins against any insect targets. Despite the identities with known toxins and the possibility that this 41.9-kDa protein may behave as a non-binary Bt toxin, it caused no mortality or symptoms of impaired growth at very high concentrations against any of the insects tested. This may be because they are outside of the insect's target range of the protein; the protein may require an unknown second component; the proteolytic activation after secretion (removal of the signal peptide) may occur in Bt but not in *E. coli*; or the protein may not, in fact, be toxic. Other species, including some from different insect orders (*i.e.*, Coleoptera and Diptera), will need to be tested to evaluate the insecticidal potential of this protein. Purified crystals from strain Hu4-2 were previously reported to be highly active against several lepidopteran species (*H. armigera*, *S. exigua*, *S. frugiperda* and *S. littoralis*), moderately toxic for two species of mosquitoes, *Culex pipiens pipiens* (Diptera: Culicidae) and *A. aegypti*, and not active against *Anopheles stephensi* (Diptera: Culicidae) [27]. Likewise, *H. armigera* and *S. littoralis* were very susceptible to the purified crystals of Leapi01 [26]. However, their complete toxin gene contents (Table 1) imply that both strains may have wider host ranges. A previous PCR screening of insecticidal toxins in strain Hu4-2 [27] detected eight *cry* gene fragments (*cryIAb*, *cryIAd*, *cryIC*, *cryID*, *cryIF*, *cryII*, *cry9E*, and *cry2Ab*), although SDS-PAGE analyses suggested that some of them are not expressed. Among the diverse crystal toxin genes found in the Hu4-2 draft genome sequence obtained in this study, we found a complete CDS for only genes *cry9Ea* and *cryIIa7*, whereas a *cry2Ab* homolog corresponded to a pseudogene with internal stop codons and multiple frame-shifted open reading frames. Sequencing may explain also why 130 to 145-kDa-crystals toxins were detected; these may be likely the product of *cry9* expression (~129.9 kDa). Similarly, the PCR screening of Leapi01 [26] detected six *cry* gene fragments (*cryIAa*, *cryIAb*, *cryICa*, *cryIDa*, *cry2Ab* and *cryIIa*) while analysis of the Leapi01 draft genome sequence, revealed seven full-length gene sequences; *cryIAa*, *cryICa*, *cryIDa*, *cry2Ab*, *cry9E* plus two *cryIIa* variants (*cryIIa7* and *cryIIa14*). These results indicated that these PCR analyses missed *cry9Ea* and misrepresented the presence of a unique *cryIIa* gene variant; additionally, although *cryIAb* was detected by PCR, annotation of the genome indicated that it was encoded by vestigial (incomplete) gene coding sequence. These previous studies [26,27] also failed to detect the 41.9-kDa proteins genes in both strains, obviously due to the lack of specific PCR primers for molecular analysis. Therefore, with the advent of NGS technology, the screening of novel or previously described holotype toxin genes has become much cheaper, faster and efficient, and may provide additional valuable information

on the structural organization and the presence of other proteins such as toxin-assisting chaperones and relevant non-coding sequences (promoters, ribosomal-binding sites, *etc*).

Acknowledgments

We acknowledge the technical support of Noelia Gorría and Itxaso Ibañez for their help with insect rearing. This research was supported by the Spanish Ministry of Science and Innovation (grant ref. AGL2009-13340-C02) and by the Universidad Pública de Navarra (PhD contract awarded to L.P.).

Authors Contribution

Leopoldo Palma and Primitivo Caballero conceived and designed the experiments. Leopoldo Palma performed the experiments. Leopoldo Palma and Primitivo Caballero analyzed the results and data. Colin Berry, Jesús Murillo and Primitivo Caballero contributed reagents, materials and funding for DNA sequencing. Leopoldo Palma, Delia Muñoz, Colin Berry, Jesús Murillo and Primitivo Caballero wrote the paper.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Schnepf, E.; Crickmore, N.; van Rie, J.; Lereclus, D.; Baum, J.; Feitelson, J.; Zeigler, D.R.; Dean, D.H. *Bacillus thuringiensis* and its pesticidal crystal proteins. *Microbiol. Mol. Biol. Rev.* **1998**, *62*, 775–806.
2. Van Frankenhuyzen, K. Insecticidal activity of *Bacillus thuringiensis* crystal proteins. *J. Invertebr. Pathol.* **2009**, *101*, 1–16.
3. Wei, J.Z.; Hale, K.; Carta, L.; Platzer, E.; Wong, C.; Fang, S.C.; Aroian, R.V. *Bacillus thuringiensis* crystal proteins that target nematodes. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 2760–2765.
4. Kondo, S.; Mizuki, E.; Akao, T.; Ohba, M. Antitrichomonal strains of *Bacillus thuringiensis*. *Parasitol. Res.* **2002**, *88*, 1090–1092.
5. Ohba, M.; Mizuki, E.; Uemori, A. Parasporin, a new anticancer protein group from *Bacillus thuringiensis*. *Anticancer Res.* **2009**, *29*, 427–433.
6. Okumura, S.; Ohba, M.; Mizuki, E.; Crickmore, N.; Côté, J.-C.; Nagamatsu, Y.; Kitada, S.; Sakai, H.; Harata, K.; Shin, T. Parasporin nomenclature. Available online: <http://parasporin.fitc.pref.fukuoka.jp/> (accessed on 03 February 2014).
7. Agaisse, H.; Lereclus, D. How does *Bacillus thuringiensis* produce so much insecticidal crystal protein? *J. Bacteriol.* **1995**, *177*, 6027–6032.
8. Manasherob, R.; Zaritsky, A.; Ben-Dov, E.; Saxena, D.; Barak, Z.; Einav, M. Effect of accessory proteins P19 and P20 on cytolytic activity of Cyt1Aa from *Bacillus thuringiensis* subsp. *israelensis* in *Escherichia coli*. *Curr. Microbiol.* **2001**, *43*, 355–364.

9. Shao, Z.; Liu, Z.; Yu, Z. Effects of the 20-kilodalton helper protein on Cry1Ac production and spore formation in *Bacillus thuringiensis*. *Appl. Environ. Microbiol.* **2001**, *67*, 5362–5369.
10. Xu, Y.; Nagai, M.; Bagdasarian, M.; Smith, T.W.; Walker, E.D. Expression of the *p20* gene from *Bacillus thuringiensis* H-14 increases Cry11A toxin production and enhances mosquito-larvicidal activity in recombinant gram-negative bacteria. *Appl. Environ. Microbiol.* **2001**, *67*, 3010–3015.
11. Estruch, J.J.; Warren, G.W.; Mullins, M.A.; Nye, G.J.; Craig, J.A.; Koziel, M.G. Vip3A, a novel *Bacillus thuringiensis* vegetative insecticidal protein with a wide spectrum of activities against lepidopteran insects. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 5389–5394.
12. Warren, G.W.; Koziel, M.G.; Mullins, M.A.; Nye, G.J.; Carr, B.; Desai, N.M.; Kostichka, K.; Duck, N.B.; Estruch, J.J. Auxiliary proteins for enhancing the insecticidal activity of pesticidal proteins. US Patent 5770696, 23 June 1998.
13. Sattar, S.; Maiti, M.K. Molecular characterization of a novel vegetative insecticidal protein from *Bacillus thuringiensis* effective against sap-sucking insect pest. *J. Microbiol. Biotechnol.* **2011**, *21*, 937–946.
14. Donovan, W.P.; Engleman, J.T.; Donovan, J.C.; Baum, J.A.; Bunkers, G.J.; Chi, D.J.; Clinton, W.P.; English, L.; Heck, G.R.; Ilagan, O.M.; *et al.* Discovery and characterization of Sip1A: A novel secreted protein from *Bacillus thuringiensis* with activity against coleopteran larvae. *Appl. Microbiol. Biotechnol.* **2006**, *72*, 713–719.
15. Berry, C. The bacterium, *Lysinibacillus sphaericus*, as an insect pathogen. *J. Invertebr. Pathol.* **2012**, *109*, 1–10.
16. Peña, G.; Miranda-Rios, J.; de la Riva, G.; Pardo-López, L.; Soberon, M.; Bravo, A. A *Bacillus thuringiensis* S-layer protein involved in toxicity against *Epilachna varivestis* (Coleoptera: Coccinellidae). *Appl. Environ. Microbiol.* **2006**, *72*, 353–360.
17. Raymond, B.; Johnston, P.R.; Nielsen-LeRoux, C.; Lereclus, D.; Crickmore, N. *Bacillus thuringiensis*: An impotent pathogen? *Trends Microbiol.* **2010**, *18*, 189–194.
18. Mac Innes, T.C.; Bouwer, G. An improved bioassay for the detection of *Bacillus thuringiensis* beta-exotoxin. *J. Invertebr. Pathol.* **2009**, *101*, 137–139.
19. Palma, L.; Hernández-Rodríguez, C.S.; Maeztu, M.; Hernández-Martínez, P.; Ruiz de Escudero, I.; Escrache, B.; Muñoz, D.; van Rie, J.; Ferré, J.; Caballero, P. Vip3C, a novel class of vegetative insecticidal proteins from *Bacillus thuringiensis*. *Appl. Environ. Microbiol.* **2012**, *78*, 7163–7165.
20. Palma, L.; Ruiz de Escudero, I.; Maeztu, M.; Caballero, P.; Muñoz, D. Screening of *vip* genes from a Spanish *Bacillus thuringiensis* collection and characterization of two novel Vip3 proteins highly-toxic to five lepidopteran crop pests. *Biol. Control* **2013**, *66*, 141–149.
21. Tan, F.; Zhu, J.; Tang, J.; Tang, X.; Wang, S.; Zheng, A.; Li, P. Cloning and characterization of two novel crystal protein genes, *cry54Aa1* and *cry30Fa1*, from *Bacillus thuringiensis* strain BtMC28. *Curr. Microbiol.* **2009**, *58*, 654–659.
22. Metzker, M.L. Sequencing technologies—The next generation. *Nat. Rev. Genet.* **2010**, *11*, 31–46.
23. Shendure, J.; Ji, H. Next-generation DNA sequencing. *Nat. Biotechnol.* **2008**, *26*, 1135–1145.
24. Iriarte, J.; Bel, Y.; Ferrandis, M.D.; Andrew, R.; Murillo, J.; Ferré, J.; Caballero, P. Environmental distribution and diversity of *Bacillus thuringiensis* in Spain. *Syst. Appl. Microbiol.* **1998**, *21*, 97–106.

25. Iriarte, J.; Porcar, M.; Lecadet, M.; Caballero, P. Isolation and characterization of *Bacillus thuringiensis* strains from aquatic environments in Spain. *Curr. Microbiol.* **2000**, *40*, 402–408.
26. Porcar, M.; Caballero, P. Molecular and insecticidal characterization of a *Bacillus thuringiensis* strain isolated during a natural epizootic. *J. Appl. Microbiol.* **2000**, *89*, 309–316.
27. Martínez, C.; Porcar, M.; López, A.; Ruiz de Escudero, I.; Pérez-Llarena, F.J.; Caballero, P. Characterization of a *Bacillus thuringiensis* strain with a broad spectrum of activity against lepidopteran insects. *Entomol. Exp. Appl.* **2004**, *111*, 71–77.
28. Altschul, S.F.; Gish, W.; Miller, W.; Myers, E.W.; Lipman, D.J. Basic local alignment search tool. *J. Mol. Biol.* **1990**, *215*, 403–410.
29. Ye, W.; Zhu, L.; Liu, Y.; Crickmore, N.; Peng, D.; Ruan, L.; Sun, M. Mining new crystal protein genes from *Bacillus thuringiensis* based on mixed plasmid-enriched genome sequencing and a computational pipeline. *Appl. Environ. Microbiol.* **2012**, *78*, 4795–4801.
30. Aziz, R.K.; Bartels, D.; Best, A.A.; DeJongh, M.; Disz, T.; Edwards, R.A.; Formsma, K.; Gerdes, S.; Glass, E.M.; Kubal, M.; *et al.* The RAST server: Rapid annotations using subsystems technology. *BMC Genomics* **2008**, *9*, 1–15.
31. Pruss, B.M.; Dietrich, R.; Nibler, B.; Martlbauer, E.; Scherer, S. The hemolytic enterotoxin HBL is broadly distributed among species of the *Bacillus cereus* group. *Appl. Environ. Microbiol.* **1999**, *65*, 5436–5442.
32. Drummond, A.J.; Ashton, B.; Buxton, S.; Cheung, M.; Cooper, A.; Duran, C.; Field, M.; Heled, J.; Kearse, M.; Markowitz, S.; *et al.* Geneious Pro v6.1.7. Available online: <http://www.geneious.com/> (accessed on 01 April 2014).
33. Sambrook, J.; Russell, D.W. *Molecular Cloning: A Laboratory Manual*, 3rd ed.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, USA, 2001.
34. Pence, R.J. The antimetabolite, imidazole as a pesticide. *J. Econ. Entomol.* **1963**, *56*, 1–7.
35. Ross, D.C.; Brown, T.M. Inhibition of larval growth in *Spodoptera frugiperda* by sublethal dietary concentrations of insecticides. *J. Agric. Food Chem.* **1982**, *30*, 193–196.
36. Bradford, M.M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **1976**, *72*, 248–254.
37. Ferré, J.; van Rie, J. Biochemistry and genetics of insect resistance to *Bacillus thuringiensis*. *Annu. Rev. Entomol.* **2002**, *47*, 501–533.
38. Sayyed, A.H.; Haward, R.; Herrero, S.; Ferré, J.; Wright, D.J. Genetic and biochemical approach for characterization of resistance to *Bacillus thuringiensis* toxin Cry1Ac in a field population of the diamondback moth, *Plutella xylostella*. *Appl. Environ. Microbiol.* **2000**, *66*, 1509–1516.
39. Oppert, B.; Kramer, K.J.; Beeman, R.W.; Johnson, D.; McGaughey, W.H. Proteinase-mediated insect resistance to *Bacillus thuringiensis* toxins. *J. Biol. Chem.* **1997**, *272*, 23473–23476.
40. Mardis, E.R. Next-generation DNA sequencing methods. *Annu. Rev. Genomics Hum. Genet.* **2008**, *9*, 387–402.
41. Zhou, X.; Ren, L.; Meng, Q.; Li, Y.; Yu, Y.; Yu, J. The next-generation sequencing technology and application. *Protein Cell* **2010**, *1*, 520–536.
42. Schuster, S.C. Next-generation sequencing transforms today’s biology. *Nat. Methods* **2008**, *5*, 16–18.

43. Baltrus, D.A.; Nishimura, M.T.; Romanchuk, A.; Chang, J.H.; Mukhtar, M.S.; Cherkis, K.; Roach, J.; Grant, S.R.; Jones, C.D.; Dangel, J.L. Dynamic evolution of pathogenicity revealed by sequencing and comparative genomics of 19 *Pseudomonas syringae* isolates. *PLoS Pathog.* **2011**, *7*, e1002132.
44. Palma, L. Universidad Pública de Navarra, Navarra, Spain. Unpublished work, 2014.
45. Porcar, M.; Iriarte, J.; Dumanoir, V.C.; Ferrandis, M.D.; Lecadet, M.M.; Ferre, J.; Caballero, P. Identification and characterization of the new *Bacillus thuringiensis* serovars pirenaica (serotype H57) and iberica (serotype H59). *J. Appl. Microbiol.* **1999**, *87*, 640–648.
46. Boetzer, M.; Henkel, C.V.; Jansen, H.J.; Butler, D.; Pirovano, W. Scaffolding pre-assembled contigs using SSPACE. *Bioinformatics* **2011**, *27*, 578–579.
47. Barton, M.D.; Barton, H.A. Scaffolder—Software for manual genome scaffolding. *Source Code Biol. Med.* **2012**, *7*, doi:10.1186/1751-0473-7-4.
48. Berry, C.; O’Neil, S.; Ben-Dov, E.; Jones, A.F.; Murphy, L.; Quail, M.A.; Holden, M.T.; Harris, D.; Zaritsky, A.; Parkhill, J. Complete sequence and organization of pBtoxis, the toxin-coding plasmid of *Bacillus thuringiensis* subsp. *israelensis*. *Appl. Environ. Microbiol.* **2002**, *68*, 5082–5095.
49. De Maagd, R.A.; Bravo, A.; Berry, C.; Crickmore, N.; Schnepf, H.E. Structure, diversity, and evolution of protein toxins from spore-forming entomopathogenic bacteria. *Annu. Rev. Genet.* **2003**, *37*, 409–433.
50. McIver, K.S.; Myles, R.L. Two DNA-binding domains of Mga are required for virulence gene activation in the group A *Streptococcus*. *Mol. Microbiol.* **2002**, *43*, 1591–1601.
51. He, J.; Wang, J.P.; Yin, W.; Shao, X.H.; Zheng, H.J.; Li, M.S.; Zhao, Y.W.; Sun, M.; Wang, S.Y.; Yu, Z.N. Complete genome sequence of *Bacillus thuringiensis* subsp. *chinensis* strain CT-43. *J. Bacteriol.* **2011**, *193*, 3407–3408.
52. McClintock, J.; Stone, T.B.; Sjoblad, R.D. A comparative review of the mammalian toxicity of *Bacillus thuringiensis*-based pesticides. *Pest Manag. Sci.* **1995**, *45*, 95–105.
53. Siegel, J.P. The mammalian safety of *Bacillus thuringiensis*-based insecticides. *J. Invertebr. Pathol.* **2001**, *77*, 13–21.
54. Han, C.S.; Xie, G.; Challacombe, J.F.; Altherr, M.R.; Bhotika, S.S.; Brown, N.; Bruce, D.; Campbell, C.S.; Campbell, M.L.; Chen, J.; *et al.* Pathogenomic sequence analysis of *Bacillus cereus* and *Bacillus thuringiensis* isolates closely related to *Bacillus anthracis*. *J. Bacteriol.* **2006**, *188*, 3382–3390.
55. Hire, R.S.; Hadapad, A.B.; Dongre, T.K.; Kumar, V. Purification and characterization of mosquitocidal *Bacillus sphaericus* BinA protein. *J. Invertebr. Pathol.* **2009**, *101*, 106–111.
56. Ruiz de Escudero, I.; Estela, A.; Porcar, M.; Martínez, C.; Oguiza, J.A.; Escriche, B.; Ferré, J.; Caballero, P. Molecular and insecticidal characterization of a CryII protein toxic to insects of the families Noctuidae, Tortricidae, Plutellidae, and Chrysomelidae. *Appl. Environ. Microbiol.* **2006**, *72*, 4796–4804.